



Fusion de données RVB-D par stéréophotométrie colorée

Quéau, Yvain; Durix, Bastien; Lucas, Tom; Boumaza, Jade; Durou, Jean-Denis; Lauze, Francois Bernard

Published in:
hal.archives-ouvertes.fr

Publication date:
2017

Document version
Også kaldet Forlagets PDF

Citation for published version (APA):

Quéau, Y., Durix, B., Lucas, T., Boumaza, J., Durou, J-D., & Lauze, F. B. (2017). Fusion de données RVB-D par stéréophotométrie colorée. *hal.archives-ouvertes.fr*, 1-25. [hal-01409663].

Fusion de données RVB-D par stéréophotométrie colorée

Yvain Quéau, Bastien Durix, Tom Lucas, Jade Boumaza, Jean-Denis Durou,
François Lauze

► To cite this version:

Yvain Quéau, Bastien Durix, Tom Lucas, Jade Boumaza, Jean-Denis Durou, et al.. Fusion de données RVB-D par stéréophotométrie colorée. 2017. <hal-01409663v2>

HAL Id: hal-01409663

<https://hal.archives-ouvertes.fr/hal-01409663v2>

Submitted on 21 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fusion de données RVB-D par stéréophotométrie colorée

Yvain Quéau¹, Bastien Durix², Tom Lucas³, Jade Boumaza⁴, Jean-Denis Durou⁵, François Lauze⁶

1. Technical University Munich, Garching, Allemagne

yvain.queau@tum.de

2. IRIT, Université de Toulouse, France

bastien.durix@enseeiht.fr

3. IRIT, Université de Toulouse, France

durou@irit.fr

4. IRIT, Université de Toulouse, France

tomlucas@live.fr

5. IRIT, Université de Toulouse, France

jade.boumaza@gmail.com

6. Department of Computer Science, University of Copenhagen, Danemark

francois@di.ku.dk

RÉSUMÉ. Nous montrons comment utiliser la stéréophotométrie colorée pour améliorer le relief fourni par un capteur RVB-D. Le capteur est équipé de trois LEDs colorées, de telle sorte que l'image RVB permet de retrouver les détails les plus fins du relief, grâce à la stéréophotométrie. Cette estimation fine du relief est fusionnée avec la carte de profondeur fournie par le capteur, grâce à une nouvelle approche différentielle et variationnelle de la stéréophotométrie adaptée aux sources ponctuelles anisotropes de type LED. Cette approche permet d'estimer la profondeur directement et de façon robuste, sans estimation préalable des normales et de l'albédo.

ABSTRACT. We improve the depth map provided by an RGB-D sensor, using color photometric stereo. The sensor is equipped with three colored LEDs, in order for thinnest geometric details to be recovered from the RGB image by photometric stereo. This thin shape estimate is then fused with the depth map provided by the sensor, through a variational approach to photometric stereo dedicated to anisotropic point light sources of LED type.

MOTS-CLÉS : reconstruction 3D, stéréophotométrie, capteurs RVB-D, méthodes variationnelles.

KEYWORDS: 3D-reconstruction, Photometric Stereo, RGB-D Sensors, Variational Methods.

Extended abstract

This paper, which unifies and extends two conference papers presented during the RFIA 2016 conference in Clermont-Ferrand (Durix *et al.*, 2016 ; Quéau, Durix *et al.*, 2016), aims at improving the accuracy of the depth map provided by an RGB-D sensor. To this end, we suggest a simple, yet very effective modification of the sensor, which consists in equipping it with three colored LEDs. The color photometric stereo technique can then be applied to the RGB data, which provides a new estimate of shape. By appropriately fusing the depth map from the sensor with that estimated using photometric stereo, the low-frequency bias in photometric stereo and the high-frequency one in the depth sensor are simultaneously eliminated.

We first describe how to model the luminous flux emitted by a LED. Considering such a light source as a nearby pointwise source, it is shown that the parameters of this model (position, orientation, and relative intensity) are easily calibrated using a standard pinhole camera and cheap additional material. The location of each LED is estimated by capturing images of multiple specular spheres (e.g., billard balls), identifying the specularities and intersecting the light rays responsible for these specularities. Given the sources locations, it is then easy to estimate their other parameters, by using images of a white planar Lambertian checkerboard and inverting Lambert’s law.

Then, we recall the variational and differential approach for photometric stereo which was presented in (Quéau, Mecca, Durou, 2016). It consists in directly estimating depth from photometric stereo images, by resorting to image ratios in order to eliminate the nonlinearities and the unknown albedo. This yields a system of quasilinear PDEs which can be solved in an approximate manner through a variational approach. At this stage, the RGB data from the proposed modified RGB-D sensor system can already be used for photometric stereo: by simultaneously illuminating the scene through the three calibrated LEDs which are colored, respectively, in red, green and blue, each channel of the RGB image can be viewed as a gray level image obtained under a different illumination. However, because our assumption that the LEDs are monochromatic does not perfectly hold, such a 3D-reconstruction remains biased.

However, this variational approach to photometric stereo is straightforward to extend in order to include a prior on the depth map. This provides us with a natural way to fuse RGB-based 3D-reconstruction by photometric stereo with the depth map provided by the sensor. We discuss a fast solution based on least-squares, and then a more robust one based on the L^1 norm. Empirical evidence that the proposed system improves the accuracy of the depth sensor is provided on several real-world datasets.

Overall, these contributions yield the first RGB-D sensor-based system which can recover high quality depth in a single shot, by appropriately combining RGB photometric stereo and depth sensing within a variational framework.

1. Introduction

La stéréophotométrie consiste à retrouver le relief d'une surface opaque à partir de $m \geq 3$ photographies prises sous le même angle, mais sous m éclairages directionnels non coplanaires (Woodham, 1980). Le cas où $m = 3$ est particulièrement intéressant, car il rend la reconstruction 3D possible avec une seule image RVB. En effet, en éclairant la surface *simultanément* par trois sources lumineuses colorées en rouge, vert et bleu, nous obtenons en une seule prise de vue l'équivalent de trois photographies prises sous trois éclairages différents, ce qui permet de reconstruire le relief de surfaces déformables en temps réel (Hernández *et al.*, 2007). Toutefois, la qualité des reconstructions 3D ainsi obtenues est médiocre (cf. figure 6-a), car avec aussi peu d'information, il est difficile de garantir une robustesse suffisante au bruit et aux données aberrantes. Or, outre l'image RVB, les capteurs RVB-D fournissent également une estimation de la distance D à la scène, ou *profondeur*, qui est elle aussi de qualité très médiocre (cf. figure 1-b). Dans cet article, qui fait la synthèse de deux articles présentés à la conférence RFIA 2016 (Durix *et al.*, 2016 ; Quéau, Durix *et al.*, 2016), nous montrons comment améliorer la qualité de la reconstruction 3D en fusionnant ces deux estimations grossières du relief.

1.1. Stéréophotométrie avec sources ponctuelles

En pratique, le caractère directionnel de l'éclairage est difficile à garantir et très peu représentatif du flux lumineux émis par les sources lumineuses à bas coût de type LED. De plus, la résolution de la stéréophotométrie avec sources ponctuelles est difficile, car l'éclairage dépend explicitement des coordonnées 3D des points de la surface, qui sont justement inconnues. La manière usuelle de résoudre ce problème consiste à estimer itérativement les normales, connaissant la profondeur, puis la profondeur, connaissant les normales (Bony *et al.*, 2013 ; Collins, Bartoli, 2012 ; Papadimitri, Favaro, 2014). Une autre possibilité consiste à estimer directement la profondeur, par déformation itérative d'un maillage représentant la surface (W. Xie *et al.*, 2015). Nous adoptons le formalisme introduit récemment dans (Quéau, Mecca, Durou, 2016), qui est à la fois variationnel et différentiel. Ce formalisme permet d'estimer directement la profondeur, de garantir une certaine robustesse aux données aberrantes et de faciliter la prise en compte de données RVB et de sources ponctuelles anisotropes. Notre montage expérimental (cf. figure 1-a) impose en effet de respecter ces contraintes.

1.2. Fusion de données RVB-D

La fusion des données issues d'un capteur de profondeur standard et d'un capteur RVB standard a fait l'objet de plusieurs études récentes. Dans (Park *et al.*, 2014 ; Yang *et al.*, 2014), les données RVB, qui possèdent une meilleure résolution et sont généralement moins bruitées, permettent d'affiner la fonction de profondeur (*depth map*). Au vu de la capacité des techniques photométriques à retrouver les détails les plus fins d'une surface, il a également été proposé de « guider » cette fusion à l'aide

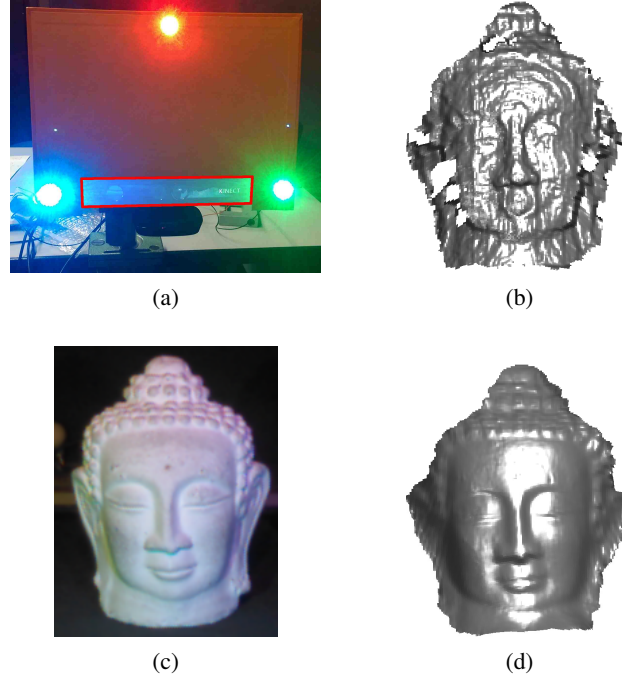


FIGURE 1. (a) Photographie de notre montage prise depuis la scène à reconstruire : le capteur RVB-D (encadré en rouge) est équipé de trois LEDs colorées en rouge, vert et bleu. (b) Rééclairage de la reconstruction 3D fournie par le capteur de profondeur, qui est bruitée et comporte des données manquantes. (c) Image RVB fournie par ce montage, qui nous permet d'appliquer la stéréophotométrie colorée. (d) Rééclairage du résultat de la fusion de données RVB-D par l'approche proposée.

d'un modèle photométrique. Dans (Or-el *et al.*, 2015), la technique du *shape-from-shading* (SfS) est utilisée en conditions non contrôlées sur les canaux RVB : la donnée de profondeur est alors vue comme un a priori permettant de limiter les ambiguïtés inhérentes au SfS. Nous pouvons espérer que le contrôle de l'éclairage permette d'améliorer la qualité des reconstructions 3D, puisqu'il rend possible l'utilisation de la stéréophotométrie, qui constitue un problème bien posé, contrairement au SfS. Nous proposons donc d'équiper le capteur RVB-D de trois LEDs colorées, afin que la stéréophotométrie colorée (Hernández *et al.*, 2007) puisse jouer le rôle de méthode « active » d'affinage du relief. En modélisant l'éclairage par un flux radial, et en proposant une méthode efficace de fusion de la profondeur D avec le relief estimé par stéréophotométrie colorée, nous obtenons un nouveau système de scannage 3D à faible coût, précis et facile à réaliser, qui est illustré sur la figure 1.

1.3. Plan de l'article

Les limites de l'approche classique de la stéréophotométrie sont rappelées dans le paragraphe 2. Dans le paragraphe 3, nous détaillons un modèle de flux lumineux émis par une LED, et décrivons une procédure d'estimation des paramètres de ce modèle. Dans le paragraphe 4, nous proposons une modélisation différentielle de la stéréophotométrie adaptée à nos conditions opératoires, dont l'inversion par une approche variationnelle est décrite dans le paragraphe 5. Enfin, nous montrons dans le paragraphe 6 comment étendre cette approche à la fusion des données d'un capteur RVB-D.

2. Approche classique de la stéréophotométrie

La stéréophotométrie se ramène généralement à un problème inverse consistant à résoudre un système d'équations qui expriment le niveau de gris en fonction du relief et de la réflectance de la surface observée, de l'éclairage et des caractéristiques de l'appareil photographique. Si l'éclairage et l'appareil photographique sont étalonnés, les seules inconnues du problème sont le relief de la surface et sa réflectance.

2.1. Inversion du modèle lambertien

Pour peu que la *mise au point* de l'appareil photographique soit correctement effectuée, il existe une bijection entre les points $\mathbf{x} \in \mathbb{R}^3$ de la partie visible d'une surface et leurs projections $\mathbf{p} \in \mathbb{R}^2$ sur le plan image. On dit d'un point \mathbf{x} et de sa projection \mathbf{p} qu'ils sont *conjugués*. La réflectance d'une surface *lambertienne* est caractérisée par l'albédo $\rho > 0$. Si les m éclairages sont supposés parallèles et uniformes, ils peuvent être caractérisés par des vecteurs $\mathbf{s}^i \in \mathbb{R}^3$, $i \in [1, m]$, orientés vers la source lumineuse et de norme proportionnelle à la densité du flux lumineux. Il est généralement supposé que l'appareil photographique fournit des images en niveaux de gris. Dans ces conditions, la stéréophotométrie se traduit comme l'estimation de l'albédo ρ et de la normale \mathbf{n} , au point \mathbf{x} de la surface conjugué du pixel \mathbf{p} , par inversion du système :

$$I^i(\mathbf{p}) = \underbrace{\rho(\mathbf{x}) \mathbf{n}(\mathbf{x})}_{\mathbf{m}(\mathbf{x})} \cdot \mathbf{s}^i, \quad i \in [1, m] \quad (1)$$

où $I^i(\mathbf{p})$ désigne le niveau de gris au pixel \mathbf{p} dans la i -ème image, $\rho(\mathbf{x})$ l'albédo en \mathbf{x} , et $\mathbf{n}(\mathbf{x})$ la normale unitaire (c'est-à-dire telle que $\|\mathbf{n}(\mathbf{x})\| = 1$) à la surface en \mathbf{x} .

Avec $m \geq 3$ éclairages *non coplanaires*, le système (1) admet une solution approchée unique $\mathbf{m}(\mathbf{x})$ en moindres carrés, qui s'exprime à l'aide de la pseudo-inverse de la matrice $[\mathbf{s}^1, \dots, \mathbf{s}^m]^\top$. De cette solution $\mathbf{m}(\mathbf{x})$, nous tirons la normale et l'albédo :

$$\mathbf{n}(\mathbf{x}) = \frac{\mathbf{m}(\mathbf{x})}{\|\mathbf{m}(\mathbf{x})\|} \quad \rho(\mathbf{x}) = \|\mathbf{m}(\mathbf{x})\| \quad (2)$$

Le champ $\mathbf{n}(\mathbf{x})$ ainsi estimé doit enfin être *intégré* en une fonction de profondeur, sachant que cette tâche recèle plus de difficultés qu'il n'y paraît (Drouot *et al.*, 2009).

2.2. Limites de l'approche classique

La variable auxiliaire $\mathbf{m} = \rho \mathbf{n}$ permet de linéariser le problème, puisque \mathbf{m} n'est pas contraint à être unitaire. Cette astuce présente néanmoins trois défauts :

- Elle n'est adaptée qu'aux images en niveaux de gris. Pour des images à plusieurs canaux, la définition de l'albédo dépend du canal de couleur considéré (Quéau, Mecca, Durou, 2016), alors que la normale n'en dépend pas. Les estimations de l'albédo et du relief sont donc généralement découplées (Barsky, Petrou, 2003), et l'intérêt de la linéarisation est perdu.
- Elle permet d'estimer non pas la fonction de profondeur, mais un champ de normales, qu'il est nécessaire d'intégrer a posteriori (Durou *et al.*, 2009). Par conséquent, toute erreur sur l'estimation des normales se propage lors de l'intégration. Cela peut être évité en ayant recours à la formulation différentielle de la stéréophotométrie, qui vise à estimer directement la profondeur (Quéau, Mecca, Durou, 2016).
- Enfin, elle est limitée à un éclairage directionnel. Pour une source ponctuelle telle qu'une LED, le vecteur \mathbf{s}^i dépend des coordonnées 3D du point \mathbf{x} de la surface éclairée. Or, ces coordonnées constituent justement les inconnues du problème. Plutôt que d'introduire la normale comme inconnue intermédiaire, la formulation différentielle s'avère plus pertinente, car elle fait directement intervenir la profondeur.

Sachant que (i) la plupart des appareils photographiques fournissent des images en couleur, et non en niveaux de gris, (ii) la finalité de la reconstruction 3D est l'estimation de coordonnées 3D, et non de normales, (iii) il est nécessaire de recourir à des montages optiques complexes pour garantir le caractère directionnel de l'éclairage, alors que les LEDs sont très peu chères et faciles à utiliser, nous proposons une approche de la stéréophotométrie qui vise à dépasser les limites de l'approche classique.

3. Modélisation du flux lumineux émis par une LED

3.1. Flux lumineux émis par une LED non colorée

Chaque LED de notre montage ayant, à l'échelle d'un visage, un volume éclairant négligeable ($\approx 1 \text{ mm}^3$), nous pouvons l'assimiler à une source ponctuelle située en un point $\mathbf{x}_s \in \mathbb{R}^3$. En tout point $\mathbf{x} \in \mathbb{R}^3$, le flux lumineux $\mathbf{s}(\mathbf{x})$ émis par une telle source est radial, c'est-à-dire colinéaire au vecteur unitaire $\mathbf{u}_r = \frac{\mathbf{x} - \mathbf{x}_s}{\|\mathbf{x} - \mathbf{x}_s\|}$. En utilisant les coordonnées sphériques (r, θ, ϕ) de \mathbf{x} , dans un repère ayant \mathbf{x}_s pour origine :

$$\mathbf{s}(\mathbf{x}) = -\frac{\Phi(\theta, \phi)}{r^2} \mathbf{u}_r \quad (3)$$

où $\Phi(\theta, \phi) \geq 0$ désigne l'intensité de la source, et où l'atténuation en $1/r^2$ découle de la conservation de l'énergie lumineuse dans un milieu non absorbant. Dans (3), le vecteur $\mathbf{s}(\mathbf{x})$ est volontairement orienté dans le sens contraire de celui de la lumière, afin de simplifier l'écriture du modèle lambertien.

Le modèle (3), qui reste très général, peut être décrit par un petit nombre de paramètres, par exemple en projetant $\Phi(\theta, \phi)$ sur la base des *harmoniques sphériques* (Basri, Jacobs, 2003). L'intensité $\Phi(\theta, \phi)$ peut également être échantillonnée au voisinage d'un plan, à l'aide d'une mire plane de réflectance connue : l'approche classique de la stéréophotométrie peut alors être utilisée, pourvu que les images de la surface à reconstruire soient « compensées » par cet échantillonnage (Angelopoulou, Petrou, 2013 ; McGunnigle, Chantler, 2003 ; Sun *et al.*, 2013), mais cela n'est justifié que si la surface à reconstruire est suffisamment proche de la mire plane. Or, il serait dommage de ne pas utiliser le fait que la plupart des LEDs émettent un flux lumineux à symétrie de révolution autour d'une direction principale repérée par un vecteur unitaire \mathbf{n}_s . Si l'angle θ est défini relativement à \mathbf{n}_s , cela signifie que $\Phi(\theta, \phi)$ est indépendant de ϕ . Le flux lumineux au point $\mathbf{x} \in \mathbb{R}^3$ provenant d'une LED située en \mathbf{x}_s s'écrit donc :

$$s(\mathbf{x}) = \frac{\Phi(\theta)}{\|\mathbf{x} - \mathbf{x}_s\|^2} \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|} \quad (4)$$

La dépendance en θ de la fonction Φ , qui est généralement décroissante sur $[0, \pi/2]$, caractérise l'*anisotropie* de la LED. Un modèle de fonction Φ vérifiant cette contrainte est celui des « sources lambertiennes imparfaites » (Bennahmias *et al.*, 2007) :

$$\Phi(\theta) = \Phi_0 \cos^\mu \theta \quad (5)$$

où $\Phi_0 = \Phi(0)$ et $\mu \geq 0$. Le modèle empirique (5), qui décrit aussi bien les sources isotropes ($\mu = 0$) que les sources lambertiennes ($\mu = 1$), a déjà été utilisé en stéréophotométrie (Mecca *et al.*, 2016 ; L. Xie *et al.*, 2015), y compris lorsque les LEDs sont disposées sur un plan parallèle au plan image, c'est-à-dire lorsque $\mathbf{n}_s = [0, 0, 1]^\top$ (Mecca *et al.*, 2014). Ce modèle a d'ailleurs fait ses preuves, puisque les constructeurs fournissent l'angle $\theta_{1/2}$ tel que $\Phi(\theta_{1/2}) = \Phi_0/2$, dont nous déduisons μ grâce à (5) :

$$\mu = -\frac{\ln(2)}{\ln(\cos \theta_{1/2})} \quad (6)$$

Pour les LEDs que nous utilisons, l'angle $\theta_{1/2}$ est égal à $\pi/3$. D'après (6), il s'ensuit que $\mu = 1$. En reportant l'expression (5) de $\Phi(\theta)$ dans (4), il vient :

$$s(\mathbf{x}) = \Phi_0 \cos^\mu \theta \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|^3} \quad (7)$$

où nous conservons la dépendance explicite en μ , afin de traiter le cas le plus général. Le modèle (7) comporte sept paramètres : trois pour les coordonnées de \mathbf{x}_s , deux pour le vecteur unitaire \mathbf{n}_s , qui intervient par l'intermédiaire de l'angle θ , et enfin Φ_0 et μ .

La *stéréophotométrie non calibrée* permet d'effectuer la reconstruction 3D d'une scène sans connaître les éclairages. Ce cas de figure a été abondamment étudié, y compris dans le cas de sources ponctuelles (Huang *et al.*, 2015 ; Papadhimetri, Favaro, 2014 ; Yeh *et al.*, 2016), mais nous avons bien sûr intérêt à étalonner les éclairages¹.

1. Il nous faut également étalonner l'appareil photographique, puisque le repère utilisé pour la reconstruction 3D lui est lié. Nous supposons que cela a été effectué préalablement à tout autre étalonnage.

3.2. Étalonnage du flux lumineux émis par une LED

La plupart des méthodes d'étalonnage d'une source ponctuelle ne tiennent compte ni de l'atténuation du flux lumineux avec l'éloignement à la source, ni de l'éventuelle anisotropie de la source, ce qui peut conduire à des résultats relativement imprécis (Ackermann *et al.*, 2013 ; Aoto *et al.*, 2012 ; Ciortan *et al.*, 2016 ; Giachetti *et al.*, 2015 ; Hara *et al.*, 2005 ; Powell *et al.*, 2001 ; Shen, Cheng, 2011 ; Takai *et al.*, 2009). Xie *et al.* montrent dans (L. Xie *et al.*, 2015) comment étalonner une LED à l'aide d'une seule mire partiellement brillante et partiellement lambertienne. Il est toutefois difficile de séparer la composante spéculaire de la composante diffuse. Nous préférons utiliser une mire brillante et une mire lambertienne. La mire brillante est utilisée pour déterminer la position de la LED par triangulation. La mire lambertienne sert à déterminer les autres paramètres par minimisation de l'erreur de reprojection, à la manière de ce qui a été proposé dans (Pintus *et al.*, 2016).

3.2.1. Mire sphérique brillante

Nous utilisons une mire sphérique brillante, en l'occurrence une boule de billard, pour déterminer la position \mathbf{x}_s de chaque LED par triangulation (Powell *et al.*, 2001). En projection perspective, le bord de la silhouette d'une sphère est une ellipse, qui peut être détectée à l'aide d'un algorithme dédié. Il est ensuite facile de déterminer la position 3D et la normale de tout point de la surface, dès lors que le rayon de la boule de billard est connu. La détection du reflet de la LED nous permet donc de déterminer, par réflexion de la ligne de visée sur la sphère, une droite de l'espace 3D passant par \mathbf{x}_s (cf. figure 2-a). En théorie, deux poses de la boule de billard suffisent à estimer \mathbf{x}_s , mais l'utilisation d'une dizaine de poses améliore la précision de cette estimation.

3.2.2. Modèle lambertien

Pour estimer la direction principale \mathbf{n}_s et l'intensité Φ_0 du modèle (7), nous utilisons une mire lambertienne. Une surface est lambertienne si la clarté apparente d'un point \mathbf{x} est indépendante de l'angle sous lequel il est observé. La *luminance* $L(\mathbf{x})$, qui est égale au flux lumineux émis par unité d'angle solide et par unité de surface apparente, est donc indépendante de la direction d'émission. Pour autant, la luminance n'est pas caractéristique de la surface, car elle dépend de l'*éclairage* $E(\mathbf{x})$, c'est-à-dire du flux lumineux reçu en \mathbf{x} par unité de surface. La relation entre luminance et éclairage s'écrit, pour une surface lambertienne :

$$L(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\pi} E(\mathbf{x}) \quad (8)$$

où l'albédo $\rho(\mathbf{x}) \in \mathbb{R}$ est une caractéristique locale de la surface égale à la proportion d'énergie lumineuse réémise : $\rho(\mathbf{x})$ vaut 1 si le point \mathbf{x} est blanc, 0 si ce point est noir.

Le paramètre $\rho(\mathbf{x})$ caractérise à lui seul la *réflectance* (ou BRDF) d'une surface lambertienne. Or, l'éclairage en un point \mathbf{x} d'une surface (lambertienne ou non) de normale $\mathbf{n}(\mathbf{x})$, éclairé par un flux lumineux $\mathbf{s}(\mathbf{x})$ et non situé dans l'ombre, s'écrit :

$$E(\mathbf{x}) = \mathbf{s}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \quad (9)$$

Après mise au point de l'appareil photographique sur un point \mathbf{x} de la scène 3D, l'éclairement $\epsilon(\mathbf{p})$ du plan image, au pixel \mathbf{p} conjugué de \mathbf{x} , est lié à la luminance $L(\mathbf{x})$ par la relation « presque linéaire » suivante (Horn, Brooks, 1989) :

$$\epsilon(\mathbf{p}) = \beta \cos^4 \alpha(\mathbf{p}) L(\mathbf{x}) \quad (10)$$

où β est un coefficient de proportionnalité caractéristique de la clarté de l'image, qui dépend de plusieurs paramètres tels que l'ouverture de l'objectif, le grandissement, etc. Quant au coefficient $\cos^4 \alpha(\mathbf{p})$, dans lequel $\alpha(\mathbf{p})$ désigne l'angle entre la ligne de visée et l'axe optique, il constitue un facteur d'assombrissement à la périphérie de l'image. Ce phénomène d'assombrissement ne doit pas être confondu avec le vignettage, car il se produirait même avec une lentille idéale (Gardner, 1947).

Avec les récepteurs photosensibles actuels, le *niveau de gris* $J(\mathbf{p})$ du pixel \mathbf{p} est quasiment proportionnel à son éclairement $\epsilon(\mathbf{p})$. En notant γ le coefficient de proportionnalité, et en combinant les égalités (8), (9) et (10), nous obtenons l'expression du niveau de gris du pixel \mathbf{p} conjugué d'un point \mathbf{x} situé sur une surface lambertienne :

$$J(\mathbf{p}) = \gamma \beta \cos^4 \alpha(\mathbf{p}) \frac{\rho(\mathbf{x})}{\pi} \mathbf{s}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \quad (11)$$

Dans (Bony *et al.*, 2013), il est conseillé d'utiliser les images au format brut pour rendre la plus effective possible cette relation de proportionnalité, qui reste malgré tout difficile à garantir. Nous reviendrons sur ce point dans le paragraphe 3.3, lorsque nous établirons l'expression des *niveaux de couleur* d'un pixel.

La bijection entre les points visibles \mathbf{x} de la surface et les pixels \mathbf{p} conjugués nous autorise à noter $\rho(\mathbf{p})$ et $\mathbf{n}(\mathbf{p})$, au lieu de $\rho(\mathbf{x})$ et $\mathbf{n}(\mathbf{x})$. Comme le facteur $\cos^4 \alpha(\mathbf{p})$ est facile à calculer en chaque pixel \mathbf{p} du récepteur photosensible, nous pouvons aisément le compenser. Nous manipulerons donc dorénavant le « niveau de gris corrigé » :

$$I(\mathbf{p}) = \frac{J(\mathbf{p})}{\cos^4 \alpha(\mathbf{p})} = \gamma \beta \frac{\rho(\mathbf{p})}{\pi} \mathbf{s}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{p}) \quad (12)$$

3.2.3. Mire plane lambertienne

Pour estimer les paramètres \mathbf{n}_s et Φ_0 du modèle (7), nous utilisons comme seconde mire un damier imprimé sur une feuille de papier blanc, elle-même collée sur un support plan (cf. figures 2-b et 2-c), en espérant que les écarts au modèle lambertien ne dégraderont pas trop la précision des estimations.

L'utilisation d'une mire plane, donc convexe, présente un avantage non négligeable : le flux lumineux $\mathbf{s}(\mathbf{x})$ en tout point \mathbf{x} de la surface est purement *primaire*, c'est-à-dire qu'il provient uniquement de la source lumineuse, sans rebond sur d'autres parties de la surface de la mire, pour peu que les murs et les objets environnants aient été recouverts de noir. Nous pouvons donc remplacer le flux lumineux $\mathbf{s}(\mathbf{x})$ de l'équation (12) par l'expression (7) du flux émis par une LED. De (7) et (12), nous déduisons le niveau de gris $I(\mathbf{p})$ de l'image d'un point \mathbf{x} de cette mire, éclairée par une LED :

$$I(\mathbf{p}) = \gamma \beta \frac{\rho(\mathbf{p})}{\pi} \Phi_0 \cos^\mu \theta \frac{(\mathbf{x}_s - \mathbf{x}) \cdot \mathbf{n}(\mathbf{p})}{\|\mathbf{x}_s - \mathbf{x}\|^3} \quad (13)$$

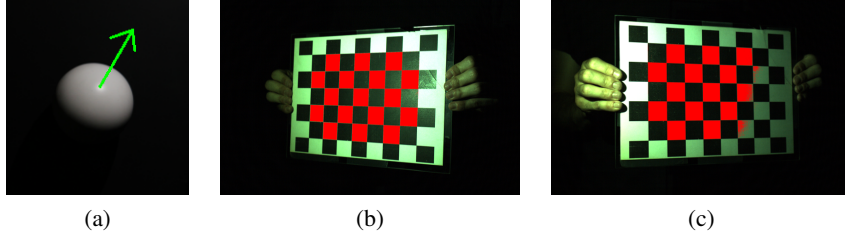


FIGURE 2. (a) La détection du reflet de la LED nous permet de déterminer, par réflexion de la ligne de visée sur la sphère, une droite de l'espace 3D passant par \mathbf{x}_s . (b)-(c) Deux des q poses de la mire plane lambertienne utilisée pour étalonner une LED. Les pixels détectés pour l'estimation de \mathbf{n}_s et Φ_0 sont indiqués en rouge.

Si $q \geq 3$ poses de la mire lambertienne sont utilisées, de nombreux algorithmes permettent d'estimer sans ambiguïté les coordonnées des points \mathbf{x}^j de sa surface, pour les différentes poses $j \in [1, q]$, les q normales \mathbf{n}^j (nous pouvons omettre la dépendance en \mathbf{p} de \mathbf{n}^j , puisque la mire est plane), ainsi que les paramètres internes de l'appareil photographique². Enfin, si l'utilisation de papier blanc ne garantit pas que $\rho(\mathbf{p}) \equiv 1$, il semble tout de même raisonnable de supposer l'albédo uniforme à l'intérieur des cases blanches, c'est-à-dire tel que $\rho(\mathbf{p}) \equiv \rho_0$. Nous regroupons l'ensemble des coefficients multiplicatifs du membre droit de (13) en un seul coefficient :

$$\Psi = \gamma \beta \frac{\rho_0}{\pi} \Phi_0 \quad (14)$$

En utilisant (14), et sachant que θ est l'angle entre les vecteurs \mathbf{n}_s et $\mathbf{x}^j - \mathbf{x}_s$, où \mathbf{x}^j est le conjugué d'un pixel \mathbf{p} de l'ensemble Ω^j des pixels blancs dans la $j^{\text{ème}}$ pose (ces pixels sont détectés automatiquement, cf. figures 2-b et 2-c), (13) s'écrit, $\forall j \in [1, q]$:

$$I^j(\mathbf{p}) = \Psi \left[\mathbf{n}_s \cdot \frac{\mathbf{x}^j - \mathbf{x}_s}{\|\mathbf{x}^j - \mathbf{x}_s\|} \right]^\mu \frac{(\mathbf{x}_s - \mathbf{x}^j) \cdot \mathbf{n}^j}{\|\mathbf{x}_s - \mathbf{x}^j\|^3} \quad (15)$$

Comme \mathbf{x}_s est déjà estimé et que la valeur de μ est fournie (indirectement) par le constructeur, les seules inconnues de l'équation (15) sont Ψ et \mathbf{n}_s . Pour les LEDs utilisées, nous savons que $\mu \neq 0$. Cela nous permet de réécrire (15) comme suit :

$$\underbrace{\Psi^{1/\mu} \mathbf{n}_s}_{\mathbf{m}_s} \cdot (\mathbf{x}^j - \mathbf{x}_s) = \left[I^j(\mathbf{p}) \frac{\|\mathbf{x}_s - \mathbf{x}^j\|^{3+\mu}}{(\mathbf{x}_s - \mathbf{x}^j) \cdot \mathbf{n}^j} \right]^{1/\mu} \quad (16)$$

L'estimation du vecteur $\mathbf{m}_s = \Psi^{1/\mu} \mathbf{n}_s$ s'écrit donc, au sens des moindres carrés :

$$\min_{\mathbf{m}_s} \sum_{j=1}^q \sum_{\mathbf{p} \in \Omega^j} \left[\mathbf{m}_s \cdot (\mathbf{x}^j - \mathbf{x}_s) - \left[I^j(\mathbf{p}) \frac{\|\mathbf{x}_s - \mathbf{x}^j\|^{3+\mu}}{(\mathbf{x}_s - \mathbf{x}^j) \cdot \mathbf{n}^j} \right]^{1/\mu} \right]^2 \quad (17)$$

2. Pour effectuer ces opérations, qui nous permettent entre autre d'étalonner l'appareil photographique, nous utilisons la boîte à outils *Computer Vision* de Matlab.

Le problème (17) peut être résolu, par exemple, à l'aide de la pseudo-inverse. De cette estimation, nous déduisons facilement les estimations des paramètres \mathbf{n}_s et Ψ :

$$\mathbf{n}_s = \frac{\mathbf{m}_s}{\|\mathbf{m}_s\|} \quad \text{et} \quad \Psi = \|\mathbf{m}_s\|^\mu \quad (18)$$

mais il est impossible de déduire de l'estimation de Ψ celle de Φ_0 , car dans la définition (14) de Ψ , le produit $\gamma \beta \frac{\rho_0}{\pi}$ est inconnu.

3.3. Flux lumineux émis par une LED colorée

Intéressons-nous maintenant au flux lumineux émis par une LED colorée. Il semble raisonnable, dans le modèle (7) de flux lumineux, de limiter la dépendance spectrale à l'intensité Φ (nous désignons par λ la longueur d'onde) :

$$\mathbf{s}(\mathbf{x}, \lambda) = \Phi(\lambda) \cos^\mu \theta \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|^3} \quad (19)$$

Ce modèle est plus complexe que (7), dans la mesure où l'intensité Φ_0 , qui est un paramètre réel, a été remplacée par le *spectre d'émission* $\Phi(\lambda)$, qui est une fonction. Plutôt que d'étalonner cette fonction, nous allons proposer un autre modèle de flux lumineux coloré, certes moins général que (19), mais adapté aux images RVB.

Pour un point \mathbf{x} d'une surface lambertienne d'albédo $\rho(\mathbf{x})$, éclairé par un flux lumineux $\mathbf{s}(\mathbf{x})$, l'expression (10) de l'éclairement $\epsilon(\mathbf{p})$ du plan image de l'appareil photographique, au pixel \mathbf{p} conjugué de \mathbf{x} , se réécrit, en utilisant (8) et (9) :

$$\epsilon(\mathbf{p}) = \beta \cos^4 \alpha(\mathbf{p}) \frac{\rho(\mathbf{x})}{\pi} \mathbf{s}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \quad (20)$$

qu'il est facile d'étendre au cas où le flux lumineux et l'albédo sont tous deux colorés :

$$\epsilon(\mathbf{p}, \lambda) = \beta \cos^4 \alpha(\mathbf{p}) \frac{\rho(\mathbf{x}, \lambda)}{\pi} \mathbf{s}(\mathbf{x}, \lambda) \cdot \mathbf{n}(\mathbf{x}) \quad (21)$$

En vertu de la bijection déjà mentionnée entre \mathbf{x} et \mathbf{p} , nous pouvons remplacer $\rho(\mathbf{x}, \lambda)$ par $\rho(\mathbf{p}, \lambda)$ et $\mathbf{n}(\mathbf{x})$ par $\mathbf{n}(\mathbf{p})$, dans l'expression (21) de $\epsilon(\mathbf{p}, \lambda)$. D'autre part, la lumière effectivement reçue par chaque photosite d'un appareil photographique traverse un filtre coloré caractérisé par son *spectre de transmission* $c_\star(\lambda)$, $\star \in \{R, V, B\}$, dont le maximum se situe, respectivement, dans le rouge, le vert ou le bleu.

La relation de proportionnalité (11), de coefficient γ , qui lie l'éclairement du récepteur photosensible à la valeur numérique enregistrée, est tout au plus valide pour un éclairage monochromatique, mais ce coefficient dépend de la longueur d'onde λ . En multipliant l'expression (21) de $\epsilon(\mathbf{p}, \lambda)$ par le produit $\gamma(\lambda) c_\star(\lambda)$, en intégrant sur l'ensemble du spectre, et en compensant le facteur $\cos^4 \alpha(\mathbf{p})$ comme en (12), nous définissons les *niveaux de couleur* (corrigés) $I_\star(\mathbf{p})$, $\star \in \{R, V, B\}$, en un pixel \mathbf{p} dont le conjugué \mathbf{x} , situé sur une surface lambertienne, est éclairé par le flux $\mathbf{s}(\mathbf{x}, \lambda)$:

$$I_\star(\mathbf{p}) = \frac{\beta}{\pi} \left[\int_{\lambda=0}^{+\infty} \gamma(\lambda) c_\star(\lambda) \rho(\mathbf{p}, \lambda) \mathbf{s}(\mathbf{x}, \lambda) d\lambda \right] \cdot \mathbf{n}(\mathbf{p}) \quad (22)$$

Nous nous restreindrons, par la suite, à la reconstruction 3D de *surfaces non colorées*, c'est-à-dire telles que $\rho(\mathbf{p}, \lambda) = \rho(\mathbf{p})$. Sous cette hypothèse, (22) s'écrit :

$$I_\star(\mathbf{p}) = \beta \frac{\rho(\mathbf{p})}{\pi} \left[\int_{\lambda=0}^{+\infty} \gamma(\lambda) c_\star(\lambda) \mathbf{s}(\mathbf{x}, \lambda) d\lambda \right] \cdot \mathbf{n}(\mathbf{p}) \quad (23)$$

Cette expression de $I_\star(\mathbf{p})$ constitue une extension de (12) aux images RVB :

$$I_\star(\mathbf{p}) = \beta \frac{\rho(\mathbf{p})}{\pi} \mathbf{s}_\star(\mathbf{x}) \cdot \mathbf{n}(\mathbf{p}) \quad (24)$$

pour peu que le vecteur $\mathbf{s}_\star(\mathbf{x})$ soit défini comme suit :

$$\mathbf{s}_\star(\mathbf{x}) = \int_{\lambda=0}^{+\infty} \gamma(\lambda) c_\star(\lambda) \mathbf{s}(\mathbf{x}, \lambda) d\lambda \quad (25)$$

En injectant l'expression (19) de $\mathbf{s}(\mathbf{x}, \lambda)$ dans (25), nous obtenons un modèle de flux lumineux coloré moins général que (19), mais mieux adapté aux images RVB :

$$\mathbf{s}_\star(\mathbf{x}) = \Phi_\star \cos^\mu \theta \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|^3} \quad (26)$$

dans lequel apparaissent les trois *intensités colorées* Φ_\star , $\star \in \{R, V, B\}$:

$$\Phi_\star = \int_{\lambda=0}^{+\infty} \gamma(\lambda) c_\star(\lambda) \Phi(\lambda) d\lambda \quad (27)$$

Le modèle (26) de flux lumineux émis par une LED colorée comporte neuf paramètres : trois pour les coordonnées de \mathbf{x}_s , deux pour le vecteur unitaire \mathbf{n}_s , plus les trois intensités colorées Φ_R , Φ_V et Φ_B et le paramètre d'anisotropie μ .

Il est notable que la définition (27) de Φ_\star dépend des fonctions $\gamma(\lambda)$ et $c_\star(\lambda)$. Les intensités colorées Φ_R , Φ_V et Φ_B ne sont donc pas des paramètres intrinsèques de la LED, puisque leur définition dépend de l'appareil photographique utilisé.

La procédure d'étalonnage du paragraphe 3.2 peut être facilement généralisée à une LED colorée. Si nous utilisons la mire plane lambertienne du paragraphe 2.2, qui est convexe, le flux lumineux incident provient uniquement de la LED. Nous pouvons donc remplacer $\mathbf{s}_\star(\mathbf{x})$ par son expression (26) dans l'expression (24) de $I_\star(\mathbf{p})$. Comme les cases blanches du damier vérifient $\rho(\mathbf{p}) \equiv \rho_0$, nous devons résoudre, pour chaque canal $\star \in \{R, V, B\}$, le problème en moindres carrés :

$$\min_{\mathbf{m}_{s,\star}} \sum_{j=1}^q \sum_{\mathbf{p} \in \Omega^j} \left[\mathbf{m}_{s,\star} \cdot (\mathbf{x}^j - \mathbf{x}_s) - \left[I_\star^j(\mathbf{p}) \frac{\|\mathbf{x}_s - \mathbf{x}^j\|^{3+\mu}}{(\mathbf{x}_s - \mathbf{x}^j) \cdot \mathbf{n}^j} \right]^{1/\mu} \right]^2 \quad (28)$$

où q désigne le nombre de poses de la mire, $\mathbf{m}_{s,\star} = \Psi_\star^{1/\mu} \mathbf{n}_s$ est défini par analogie avec \mathbf{m}_s , et $\Psi_\star = \beta \frac{\rho_0}{\pi} \Phi_\star$. La résolution de chaque problème (28) nous permet d'estimer une intensité colorée Φ_R , Φ_V ou Φ_B , au facteur $\beta \frac{\rho_0}{\pi}$ près. Nous savons donc estimer les vecteurs $\bar{\mathbf{s}}_\star(\mathbf{x}) = \beta \frac{\rho_0}{\pi} \mathbf{s}_\star(\mathbf{x})$. D'après (24) :

$$I_\star(\mathbf{p}) = \frac{\rho(\mathbf{p})}{\rho_0} \bar{\mathbf{s}}_\star(\mathbf{x}) \cdot \mathbf{n}(\mathbf{p}) \quad (29)$$

4. Un nouveau modèle différentiel pour la stéréophotométrie

L'approche classique de la stéréophotométrie, qui a été rappelée dans le paragraphe 2, est adaptée aux éclairages directionnels et aux images en niveaux de gris. Nous décrivons ci-après un nouveau modèle de stéréophotométrie qui tient compte du caractère radial de l'éclairage émis par une LED à spectre coloré.

4.1. Modèle photométrique différentiel

Afin d'obtenir directement une fonction de profondeur $z : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ caractérisant le relief de la surface photographiée, nous choisissons la formulation différentielle de la stéréophotométrie. Pour ce faire, il nous faut préciser le lien entre la normale à la surface et le gradient de la fonction de profondeur.

Nous utilisons le modèle d'appareil photographique en « trou d'épingle », caractérisé par la *matrice des paramètres internes* :

$$\mathbf{K} = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (30)$$

où $\mathbf{p}_0 = [u_0, v_0]^\top$ est le *point principal*, et les paramètres f_u et f_v , exprimés en pixels (px), représentent le produit de la distance focale (en mm) par les dimensions d'un pixel dans les directions horizontale et verticale (en $px.mm^{-1}$). Ces paramètres sont supposés connus. La reconstruction 3D consiste à estimer :

$$\mathbf{x} = z \left[\frac{u - u_0}{f_u}, \frac{v - v_0}{f_v}, 1 \right]^\top \quad (31)$$

c'est-à-dire à déterminer la valeur de z associée à chaque pixel. Nous notons :

$$\begin{aligned} z : \quad \Omega \subset \mathbb{R}^2 &\rightarrow \mathbb{R} \\ \mathbf{p} = [u, v]^\top &\mapsto z(\mathbf{p}) \end{aligned} \quad (32)$$

la fonction de profondeur associant à tout pixel \mathbf{p} la valeur de z correspondante, et $\nabla z(\mathbf{p}) = [\partial_u z(\mathbf{p}), \partial_v z(\mathbf{p})]^\top$ le gradient de cette fonction en \mathbf{p} .

Il est facile de montrer que la normale $\mathbf{n}(\mathbf{x})$ s'écrit :

$$\mathbf{n}(\mathbf{x}) = \frac{1}{d_z(\mathbf{p})} \begin{bmatrix} f_u \partial_u z(\mathbf{p}) \\ f_v \partial_v z(\mathbf{p}) \\ \underbrace{-z(\mathbf{p}) - [\mathbf{p} - \mathbf{p}_0]^\top \cdot \nabla z(\mathbf{p})}_{\bar{\mathbf{n}}(\mathbf{p})} \end{bmatrix} \quad (33)$$

où le dénominateur $d_z(\mathbf{p}) = \|\bar{\mathbf{n}}(\mathbf{p})\|$, dont la dépendance en z est non linéaire, permet de rendre la normale unitaire.

En remplaçant, dans (29), la normale $\mathbf{n}(\mathbf{x})$ par l'expression différentielle (33), et en développant le produit scalaire, nous obtenons, pour chaque canal $\star \in \{R, V, B\}$ de chaque image $i \in [1, m]$:

$$I_{\star}^i(\mathbf{p}) = \frac{\rho(\mathbf{p})}{\rho_0 d_z(\mathbf{p})} [\mathbf{t}_{\star}^i(\mathbf{x}) \cdot \nabla z(\mathbf{p}) - \bar{s}_{\star,3}^i z(\mathbf{p})] \quad (34)$$

où :

$$\mathbf{t}_{\star}^i(\mathbf{x}) = \begin{bmatrix} f_u \bar{s}_{\star,1}^i(\mathbf{x}) - (u - u_0) \bar{s}_{\star,3}^i(\mathbf{x}) \\ f_v \bar{s}_{\star,2}^i(\mathbf{x}) - (v - v_0) \bar{s}_{\star,3}^i(\mathbf{x}) \end{bmatrix} \quad (35)$$

et où $\bar{s}_{\star,1}^i(\mathbf{x})$, $\bar{s}_{\star,2}^i(\mathbf{x})$ et $\bar{s}_{\star,3}^i(\mathbf{x})$ désignent les trois composantes du vecteur $\bar{\mathbf{s}}_{\star}^i(\mathbf{x})$.

Le système d'équations (34), lorsque $i \in [1, m]$, est un système d'EDPs non linéaires en z . Il présente l'avantage appréciable de faire intervenir explicitement la fonction de profondeur z , dont la connaissance est requise pour calculer l'éclairage incident en un point \mathbf{x} de la surface. Par rapport à (24), cette formulation présente cependant deux difficultés : elle est *non linéaire*, et ne peut être résolue localement, c'est-à-dire en un point \mathbf{x} indépendamment de ses voisins.

Nous montrons dans le prochain paragraphe comment éliminer les albédos ρ_0 et $\rho(\mathbf{p})$ de (34), ainsi qu'une partie de la non linéarité de ces équations, afin d'obtenir un système d'EDPs *quasi-linéaires* en z .

4.2. Élimination partielle de la non linéarité

La première difficulté du système d'EDPs (34) est la non linéarité due au dénominateur $d_z(\mathbf{p})$, dont le rôle est de rendre la normale unitaire. L'absence de connaissance a priori sur les albédos ρ_0 et $\rho(\mathbf{p})$ constitue une autre difficulté. Toutefois, ces différents facteurs sont indépendants du numéro i de l'éclairage. À partir d'une paire d'équations (34) correspondant au même canal $\star \in \{R, V, B\}$, mais provenant de deux images distinctes (il existe $c = C_m^2$ paires d'images de ce type, d'indices i_1 et i_2 tels que $1 \leq i_1 < i_2 \leq m$), nous pouvons donc déduire les égalités suivantes :

$$\frac{\rho(\mathbf{p})}{\rho_0 d_z(\mathbf{p})} = \frac{I_{\star}^{i_1}(\mathbf{p})}{\mathbf{t}_{\star}^{i_1}(\mathbf{x}) \cdot \nabla z(\mathbf{p}) - \bar{s}_{\star,3}^{i_1} z(\mathbf{p})} = \frac{I_{\star}^{i_2}(\mathbf{p})}{\mathbf{t}_{\star}^{i_2}(\mathbf{x}) \cdot \nabla z(\mathbf{p}) - \bar{s}_{\star,3}^{i_2} z(\mathbf{p})} \quad (36)$$

Des identifications similaires ont déjà été suggérées pour la stéréophotométrie. Il a notamment été montré dans (Quéau, Mecca, Durou, 2016) qu'elles permettaient de résoudre le problème de la stéréophotométrie pour une surface colorée éclairée par des sources blanches.

Nous tirons des égalités (36) l'EDP en z suivante :

$$\underbrace{[I_{\star}^{i_1}(\mathbf{p}) \mathbf{t}_{\star}^{i_2}(\mathbf{x}) - I_{\star}^{i_2}(\mathbf{p}) \mathbf{t}_{\star}^{i_1}(\mathbf{x})] \cdot \nabla z(\mathbf{p})}_{\mathbf{a}_z^{i_1, i_2}(\mathbf{p}) \in \mathbb{R}^2} = \underbrace{[I_{\star}^{i_1}(\mathbf{p}) \bar{s}_{\star,3}^{i_2}(\mathbf{x}) - I_{\star}^{i_2}(\mathbf{p}) \bar{s}_{\star,3}^{i_1}(\mathbf{x})] z(\mathbf{p})}_{b_z^{i_1, i_2}(\mathbf{p}) \in \mathbb{R}} \quad (37)$$

Les champs vectoriels \bar{s} et \mathbf{t} dépendent de z , car $\bar{s}(\mathbf{x}) = \bar{s}(\mathbf{x}(z))$ et $\mathbf{t}(\mathbf{x}) = \mathbf{t}(\mathbf{x}(z))$, mais ne dépendent pas de ∇z : l'EDP (37) est donc *quasi-linéaire*.

En formant le champ matriciel $\mathbf{A}_z : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^{3c \times 2}$, où $c = C_m^2$, et le champ vectoriel $\mathbf{b}_z : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^{3c}$, où chaque ligne est constituée, respectivement, de l'un des vecteurs lignes $\mathbf{a}_z^{i_1, i_2}(\mathbf{p})^\top$, et de l'un des scalaires $b_z^{i_1, i_2}(\mathbf{p})$, nous obtenons le système d'EDPs quasi-linéaires en z suivant :

$$\mathbf{A}_z(\mathbf{p}) \nabla z(\mathbf{p}) = \mathbf{b}_z(\mathbf{p}) \quad (38)$$

Cette formulation différentielle est plus simple que la formulation différentielle originale (34), dans la mesure où la non linéarité la plus importante, due au dénominateur d_z , a été éliminée. Il nous reste à proposer une méthode de résolution tenant compte du caractère redondant des EDPs du système (38), qui sont en général incompatibles à cause des écarts au modèle (de tels écarts sont inévitables avec des données réelles). Cela fait l'objet du prochain paragraphe.

5. Résolution variationnelle du système d'EDPs (38)

La résolution *exacte* du système (38) n'est pas souhaitable, à cause des écarts au modèle lambertien. Elle est d'ailleurs impossible, vu que ces équations sont généralement incompatibles, à cause justement des écarts au modèle, sauf dans le cas où nous ne disposerions que de $m = 2$ images en niveaux de gris (Quéau, Mecca *et al.*, 2017). Le formalisme variationnel permet de prendre en compte à la fois la redondance d'information et les écarts au modèle. Nous allons commencer par résoudre ce système au sens d'un critère en moindres carrés. En faisant l'hypothèse que chaque égalité (38) n'est vérifiée qu'à un bruit additif près, et que ce bruit est gaussien et de moyenne nulle, la solution optimale (au sens du maximum de vraisemblance) est obtenue en minimisant la somme des carrés des résidus de (38) sur l'ensemble du domaine Ω :

$$\min_z \|\mathbf{A}_z \nabla z - \mathbf{b}_z\|_{L^2(\Omega)}^2 \quad (39)$$

En réalité, toute hypothèse sur le bruit devrait être formulée sur les images, et non sur le modèle (38), qui a été obtenu à partir des quotients (36) : si le bruit sur les images est gaussien, alors le bruit sur ces quotients suit une loi de Cauchy (Hinkley, 1969). Par conséquent, la solution en moindres carrés (39) est bien le meilleur estimateur linéaire non biaisé, mais il ne s'agit pas pour autant de la solution optimale (au sens du maximum de vraisemblance), qui nécessiterait de recourir à un M-estimateur non convexe. Cette piste, qui a été suivie récemment dans (Quéau, Wu *et al.*, 2017), est toutefois laissée en perspective.

La non linéarité du problème en moindres carrés (39) provient de la dépendance en z des champs \mathbf{A}_z et \mathbf{b}_z . D'après (37), (35) et (26), cette dépendance est due à l'atténuation du flux lumineux en $1/\|\mathbf{x}_s - \mathbf{x}\|^2$, qui caractérise les sources ponctuelles, et à son atténuation en $\cos^\mu \theta$, qui caractérise les sources anisotropes. Si l'objet observé est suffisamment « petit » par rapport à sa distance aux sources lumineuses, alors la variation du résidu $\|\mathbf{A}_z \nabla z - \mathbf{b}_z\|_{L^2(\Omega)}^2$ est surtout due au gradient ∇z .

Ce raisonnement nous invite, malgré l'absence de garantie sur la convexité du problème (39), à nous orienter vers un schéma numérique *de type point fixe*, qui découple l'estimation de z des estimations de \mathbf{A}_z et de \mathbf{b}_z . Commençons par réécrire (39) sous la forme d'un problème contraint :

$$\begin{aligned} \min_z \quad & \|\mathbf{A}\nabla z - \mathbf{b}\|_{L^2(\Omega)}^2 \\ \text{s.c.} \quad & \begin{bmatrix} \mathbf{A} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_z \\ \mathbf{b}_z \end{bmatrix} \end{aligned} \quad (40)$$

Une itération du schéma de résolution que nous proposons, de type point fixe, s'écrit :

$$z^{(k+1)} = \underset{z}{\operatorname{argmin}} \|\mathbf{A}^{(k)}\nabla z - \mathbf{b}^{(k)}\|_{L^2(\Omega)}^2 \quad (41)$$

$$\begin{bmatrix} \mathbf{A}^{(k+1)} \\ \mathbf{b}^{(k+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{z^{(k+1)}} \\ \mathbf{b}_{z^{(k+1)}} \end{bmatrix} \quad (42)$$

Le problème (41) est un problème variationnel en moindres carrés linéaires. En discrétisant l'opérateur gradient par différences finies (dans l'implémentation actuelle, nous utilisons des différences finies avant d'ordre 1, avec une condition au bord de type Neumann), nous transformons (41) en un problème d'optimisation discrète, dont la condition nécessaire d'optimalité est tout simplement un système linéaire ayant pour inconnues les valeurs de z :

$$\underbrace{(\overline{\mathbf{A}\mathbf{D}})^\top \overline{\mathbf{A}\mathbf{D}}}_{\overline{\overline{\mathbf{A}}}} \mathbf{z} = (\overline{\mathbf{A}\mathbf{D}})^\top \overline{\mathbf{b}} \quad (43)$$

où $\overline{\mathbf{A}}$, $\overline{\mathbf{b}}$ et \mathbf{D} désignent, respectivement, les discrétisations du champ matriciel $\mathbf{A}^{(k)}$, du champ vectoriel $\mathbf{b}^{(k)}$ et de l'opérateur gradient, et où le vecteur colonne \mathbf{z} est construit par concaténation des valeurs de z .

La matrice $\overline{\overline{\mathbf{A}}}$ du système linéaire (43) est symétrique, *semi-définie positive*. Son noyau est de dimension 1. La déficience de rang de $\overline{\overline{\mathbf{A}}}$ provient de ce que, pour tout réel k , $\mathbf{A}\nabla z - \mathbf{b} = \mathbf{A}\nabla(z + k) - \mathbf{b}$. Elle ne serait pas survenue avec le problème en moindres carrés non linéaires initial (39), car les champs \mathbf{A}_z et \mathbf{b}_z ne sont pas invariants par translation le long de l'axe des z . Néanmoins, comme le système (43) est issu d'un problème en moindres carrés, nous pouvons le résoudre localement par l'algorithme du gradient conjugué (à l'itération $(k + 1)$ du schéma de résolution, la solution précédente $\mathbf{z}^{(k)}$ est choisie comme initialisation).

Cet algorithme permet de résoudre le problème de la stéréophotométrie avec sources ponctuelles. Nous montrons sur la figure 3 un résultat de reconstruction 3D obtenu à partir de $m = 8$ images RVB d'une statuette en plâtre, éclairée par $m = 8$ LEDs étalonnées selon la procédure décrite dans le paragraphe 3.

Malgré tout, le temps d'acquisition des $m = 8$ images n'est pas négligeable. Pour atteindre l'objectif d'un traitement en temps réel, il nous faut réduire le nombre d'images utilisées, au risque de compromettre la robustesse de la méthode.



FIGURE 3. *Reconstruction 3D obtenue par stéréophotométrie avec sources ponctuelles. En haut : trois (parmi $m = 8$) images RVB d'une statuette en plâtre, éclairée par $m = 8$ LEDs. En bas : résultat de l'approche variationnelle proposée.*

Il est possible d'appliquer notre méthode à partir d'une seule image RVB, en utilisant une LED verte, une LED rouge et une LED bleue. Chaque canal coloré de l'image RVB peut alors être vu comme une image en niveaux de gris, prise sous l'éclairage d'une seule LED (celle dont la couleur correspond au canal considéré). La figure 4 montre toutefois que la qualité des reconstructions 3D obtenues ainsi est très inférieure à celle des résultats obtenus avec un grand nombre d'images. Ceci vient probablement du fait que les LEDs que nous utilisons ne sont pas parfaitement monochromatiques.

Pour pouvoir appliquer la stéréophotométrie dans le cas où les sources sont ponctuelles et où une seule image RVB est utilisée, il nous faudrait réaliser un étalonnage précis du spectre d'émission de chaque LED. Dans le prochain paragraphe, nous proposons une autre piste, qui consiste à étendre notre approche à la fusion de deux estimations imprécises de la profondeur, l'une provenant de la stéréophotométrie à une seule image RVB, l'autre d'un capteur de profondeur. La qualité des reconstructions 3D obtenues par ce procédé sera comparable à celle des reconstructions 3D effectuées par stéréophotométrie avec un grand nombre d'images.

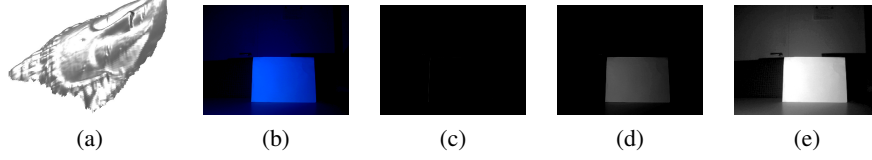


FIGURE 4. (a) *Reconstruction 3D par stéréophotométrie obtenue avec le montage de la figure 1-a, à partir d'une seule image RVB.* (b) *Photographie couleur d'une feuille de papier blanc prise sous l'éclairage de la LED bleue.* (c)-(d)-(e) *Décomposition de cette photographie dans les trois canaux. Le canal vert montre que la LED bleue est loin d'être parfaitement monochromatique. Plutôt que d'étalonner précisément le spectre de chaque LED colorée, afin de corriger le biais de la reconstruction 3D (a), l'originalité de notre approche consiste à fusionner cette reconstruction 3D avec une autre estimation grossière du relief fournie par le capteur RVB-D, afin d'obtenir une reconstruction 3D de qualité comparable à celle de la figure 3.*

6. Fusion de données RVB-D

Nous nous intéressons maintenant aux capteurs de type RVB-D. La donnée de profondeur fournie par ces capteurs est généralement très bruitée. Pour en améliorer la qualité, nous proposons une solution facile à réaliser et très peu coûteuse, qui consiste à adjoindre au capteur RVB-D trois sources lumineuses colorées de type LED (cf. figure 1-a). L'approche variationnelle décrite dans la partie 5 nous permettra de fusionner la reconstruction 3D obtenue par stéréophotométrie à une seule image RVB (Hernández *et al.*, 2007) aux données fournies par le capteur de profondeur.

6.1. Modèle variationnel L^2 - L^2

Soit z_0 la fonction de profondeur fournie par le capteur RVB-D, qui est généralement bruitée et incomplète (données de profondeur manquantes). Nous pourrions chercher à débruiter et à compléter z_0 par une approche variationnelle, en minimisant la somme d'un terme d'attache aux données de la forme $\lambda \|z - z_0\|_{L^2(\Omega)}^2$, $\lambda > 0$, et d'un terme de régularisation spatiale de la forme $\|\nabla z\|_{L^2(\Omega)}^2$ (Sobolev) ou $\|\nabla z\|_{L^2(\Omega)}$ (variation totale). Cependant, un tel modèle ne tiendrait pas compte du fait que nous disposons, en plus de la donnée de profondeur z_0 , d'une image RVB.

Or, une image RVB suffit à appliquer la *stéréophotométrie colorée* : chaque canal R, V ou B peut être interprété comme une image en niveaux de gris, obtenue sous un éclairage que nous supposons radial et uniquement dû à la LED de même couleur. En chaque pixel \mathbf{p} , les trois niveaux de couleur nous permettent donc d'écrire un système de $c = C_3^2 = 3$ EDPs de type (37). Nous suggérons donc d'utiliser l'équation (38), où $\mathbf{A}_z(\mathbf{p}) \in \mathbb{R}^{3 \times 2}$ et $\mathbf{b}_z(\mathbf{p}) \in \mathbb{R}^3$, comme terme de régularisation quadratique d'un

problème variationnel, dans lequel la connaissance a priori z_0 est prise en compte par un terme d'attache aux données également quadratique :

$$\min_z \left\{ \underbrace{\|\mathbf{A}_z \nabla z - \mathbf{b}_z\|_{L^2(\Omega)}^2 + \|\sqrt{\lambda}(z - z_0)\|_{L^2(\Omega)}^2}_{\mathcal{E}(z)} \right\} \quad (44)$$

La complétion des données de profondeur manquantes est obtenue en attribuant à λ une dépendance spatiale : nous fixons $\lambda(\mathbf{p}) = \lambda_0 > 0$ là où une donnée de profondeur $z_0(\mathbf{p})$ est disponible, et $\lambda(\mathbf{p}) = 0$ partout ailleurs. Dans (44), le terme de régularisation joue donc à la fois un rôle de débruitage et d'interpolation (*inpainting*).

Contrairement aux modèles de débruitage usuels, le modèle (44) ne comporte pas de terme explicite de régularisation spatiale. Dans le cas du *shape-from-shading*, qui est un problème mal posé, un terme de ce type est nécessaire, mais cela n'est plus le cas pour la stéréophotométrie calibrée, qui constitue un problème bien posé.

La résolution du problème (44) ne diffère guère de celle du problème (39) : l'ajout de l'a priori modifie seulement les coefficients diagonaux de la matrice $\overline{\mathbf{A}}$ définie en (43). En outre, comme ce terme n'est pas invariant par translation, le rang de la matrice $\overline{\mathbf{A}}$ devient maximal dès lors que $\lambda(\mathbf{p}) > 0$ en au moins un pixel $\mathbf{p} \in \Omega$.

La figure 5 montre l'évolution du relief estimé en fonction des itérations, avec les données de la figure 1. La figure 6 montre que le réglage de la valeur du paramètre λ_0 revient à choisir un « compromis » entre la donnée de profondeur z_0 , qui est bruitée, et la reconstruction 3D par stéréophotométrie qui, à cause du faible nombre $m = 3$ d'images, est assez peu fiable elle aussi. Enfin, la figure 7 permet de comparer notre méthode à celle d'Or-El et al. (Or-el *et al.*, 2015).

6.2. Modèle variationnel robuste L^1 - L^1

L'approche en moindres carrés, bien que très rapide, n'est pas robuste aux données aberrantes. Par exemple, en présence d'ombres portées, d'ombres propres ou de reflets brillants, le niveau de couleur mesuré n'est pas fidèle au modèle (29). D'autre part, la profondeur fournie par le capteur RVB-D est erronée en certains points, à cause du recalage de l'image RVB avec la fonction de profondeur, qui est manifestement imprécis. Nous proposons de traiter ces deux types de données aberrantes (photométriques et de profondeur) comme des écarts *parcimonieux*, qui peuvent être pris en compte en remplaçant, dans le problème (44), la norme L^2 au carré par la norme L^1 :

$$\min_z \left\{ \|\mathbf{A}_z \nabla z - \mathbf{b}_z\|_{L^1(\Omega)} + \|\sqrt{\lambda}(z - z_0)\|_{L^1(\Omega)} \right\} \quad (45)$$

Le problème (45) est nettement plus difficile à résoudre que son équivalent quadratique (44). Cependant, l'introduction d'opérateurs proximaux permet de le décompo-

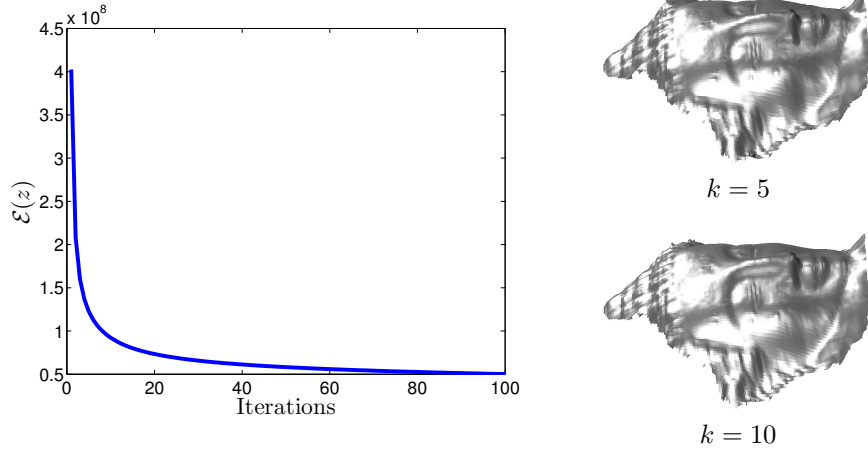


FIGURE 5. À gauche : graphe de l'énergie $\mathcal{E}(z)$ définie en (44), en fonction des itérations, pour $\lambda_0 = 10^4$. Malgré l'absence de garantie sur la convergence du schéma de résolution proposé, nous observons expérimentalement que $\mathcal{E}(z)$ décroît au cours des itérations. À droite : relief estimé après k itérations. Même si la convergence du schéma vers la solution de (44) est plutôt lente, 5 à 10 itérations (≈ 2 s de calcul sur un processeur I7 récent, avec des codes Matlab non optimisés) suffisent en pratique à obtenir un résultat qualitativement satisfaisant.

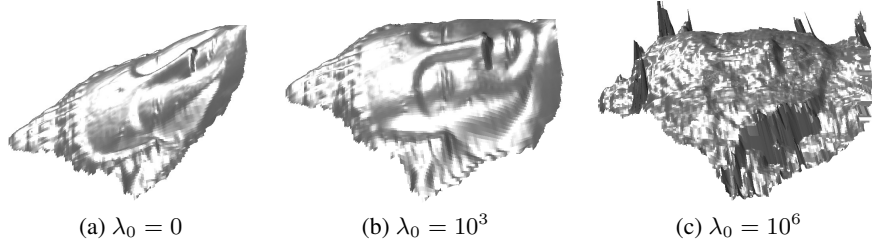


FIGURE 6. Reconstructions 3D obtenues pour différentes valeurs du paramètre λ_0 . (a) Contrairement aux cas où $m > 3$ images RVB sont disponibles (cf. figure 3), la stéréophotométrie à une seule image RVB comporte un biais. (b) Ce biais disparaît lorsque λ_0 augmente. (c) Toutefois, le relief obtenu avec une valeur trop élevée de λ_0 est similaire au relief très bruité fourni par le capteur RVB-D (cf. figure 7-a).

ser en une séquence de problèmes quadratiques plus faciles à résoudre. En introduisant une variable auxiliaire \bar{z} , nous réécrivons (45) sous la forme suivante :

$$\min_{z, \bar{z}} \left\{ \|\mathbf{A}_z \nabla z - \mathbf{b}_z\|_{L^1(\Omega)} + \frac{1}{2t} \|z - \bar{z}\|_{L^2(\Omega)}^2 + \|\sqrt{\lambda}(\bar{z} - z_0)\|_{L^1(\Omega)} \right\} \quad (46)$$

où t est un paramètre fixé par l'utilisateur (dans les tests, nous choisissons $t = \frac{1}{2\lambda_0}$).

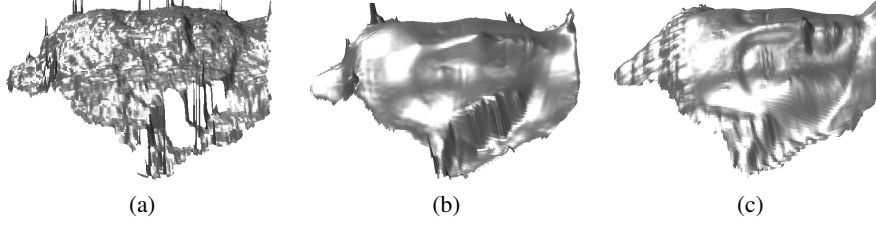


FIGURE 7. (a) Relief z_0 fourni par le capteur RVB-D. (b) Fusion de données RVB-D par shape-from-shading non calibré (Or-el et al., 2015). (c) Reconstruction 3D L^2 - L^2 obtenue pour $\lambda = 10^4$, valeur qui donne visuellement le meilleur résultat.

Nous résolvons le problème (46) par un schéma d'optimisation alternée. L'estimation de z , à \bar{z} fixé, se ramène au problème (44) par un schéma ADMM (Boyd et al., 2011). L'estimation de \bar{z} , à z fixé, admet la solution explicite suivante :

$$\bar{z} = \begin{cases} z^{(k+1)} + \frac{\lambda_0}{2t} & \text{si } 2t(z - z_0) \leq -\lambda_0 \\ z^{(k+1)} - \frac{\lambda_0}{2t} & \text{si } 2t(z - z_0) \geq \lambda_0 \\ z_0 & \text{sinon} \end{cases} \quad (47)$$

La figure 8 montre que le remplacement de la norme L^2 au carré par la norme L^1 permet bien d'améliorer les résultats, en particulier au niveau des discontinuités de profondeur. Nous fixons $\lambda_0 = 1$ pour la reconstruction L^1 - L^1 , ayant observé expérimentalement que λ_0 devait être plus faible que pour l'approche L^2 - L^2 . Bien que le terme $\|\mathbf{A}_z \nabla z - \mathbf{b}_z\|_{L^1(\Omega)}$ soit à la fois non lisse et non convexe, ce qui ne permet pas de garantir la convergence de l'algorithme ADMM, nous n'avons pas constaté d'instabilité particulière. La convergence est atteinte en une dizaine d'itérations, ce qui nécessite environ 1 minute de calcul dans notre environnement. Ce temps de calcul pourrait probablement être diminué en recourant à des techniques de parallélisation massive sur GPU.

7. Conclusion et perspectives

Nous avons proposé une reformulation de la stéréophotométrie sous la forme d'un système d'EDPs quasi-linéaires, et une approche variationnelle permettant de résoudre ce système. Nous avons également montré comment étendre cette approche à la prise en compte d'une connaissance a priori sur la profondeur, fournie par exemple par un capteur RVB-D. Enfin, nous avons montré expérimentalement l'intérêt de notre approche, grâce à une modification très simple des capteurs RVB-D : en les équipant de trois LEDs colorées en rouge, vert et bleu, la qualité du relief fourni peut être améliorée de façon significative. Même s'il nous reste à quantifier cette amélioration, en comparant une reconstruction 3D à la vérité terrain, obtenue par exemple par un scanner à laser, les résultats sont qualitativement très prometteurs.

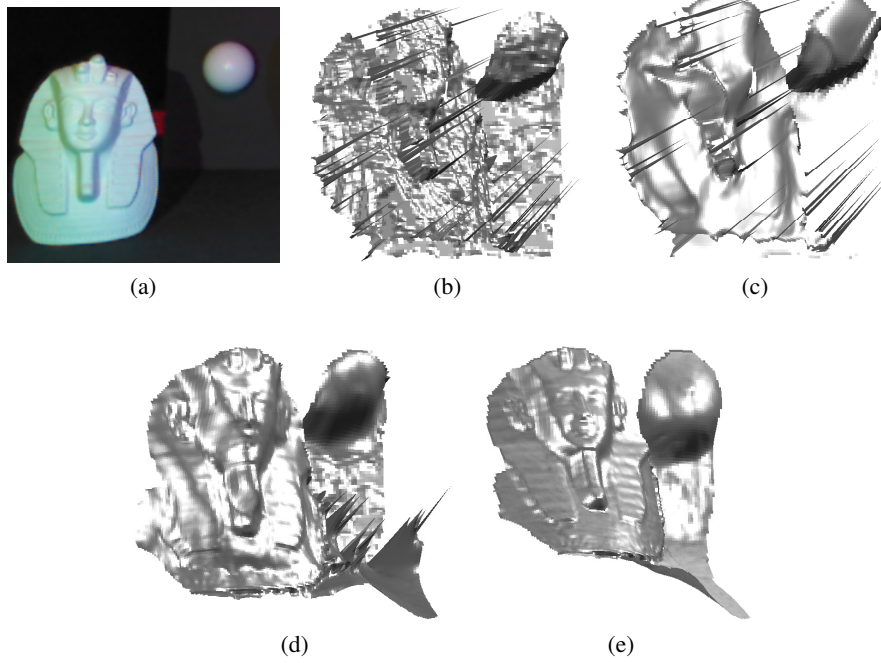


FIGURE 8. (a) Exemple d'image RVB obtenue avec notre montage (cf. figure 1-a). (b) Fonction de profondeur z_0 associée. (c) Fusion de données RVB-D par *shape-from-shading* non calibré (Or-el et al., 2015) : le relief est sur-lissé. (d) Fusion de données RVB-D par l'approche L^2 - L^2 : les discontinuités de profondeur induisent des artefacts, (e) qui sont fortement atténués en utilisant l'approche robuste L^1 - L^1 .

Notre approche repose cependant sur l'hypothèse que l'albédo de la surface observée est indépendant du canal, qui nous a permis de simplifier l'expression (22) du niveau de couleur $I_\star(\mathbf{p})$ sous la forme (24). En pratique, cette hypothèse limite l'intérêt de notre méthode. Une extension naturelle de notre travail consistera à nous passer de cette hypothèse, en recourant à l'utilisation d'images multi-, voire hyper-spectrales.

Enfin, nous avons supposé que les sources lumineuses étaient parfaitement monochromatiques, afin de garantir que l'éclairage dans le canal \star n'était dû qu'à la source lumineuse de couleur \star . Or, cette hypothèse n'est pas parfaitement vérifiée. La figure 4 montre que nous avons négligé certains niveaux de couleur qui, quoique faibles, ne sont pas nuls. Pour améliorer les résultats, nous pourrions donc étalonner le spectre d'émission de chaque LED, ainsi que le spectre de transmission de chaque type de filtre. Ceci nous permettrait sans doute de corriger, au moins en partie, le biais du résultat obtenu par stéréophotométrie à une seule image RVB (cf. figure 6-a).

Bibliographie

- Ackermann J., Fuhrmann S., Goesele M. (2013). Geometric Point Light Source Calibration. In *Proceedings of the 18th International Workshop on Vision, Modeling & Visualization*, p. 161–168. Lugano, Suisse.
- Angelopoulou M.-E., Petrou M. (2013). Uncalibrated flatfielding and illumination vector estimation for photometric stereo face reconstruction. *Machine Vision and Applications*, vol. 25, n° 5, p. 1317–1332.
- Aoto T., Taketomi T., Sato T., Mukaigawa Y., Yokoya N. (2012). Position estimation of near point light sources using a clear hollow sphere. In *Proceedings of the 21st International Conference on Pattern Recognition*, p. 3721–3724. Tsukuba, Japon.
- Barsky S., Petrou M. (2003). The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, n° 10, p. 1239–1252.
- Basri R., Jacobs D. W. (2003). Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, n° 2, p. 218–233.
- Bennahmias M., Arik E., Yu K., Voloshenko D., Chua K., Pradhan R. *et al.* (2007). Modeling of non-Lambertian sources in lighting applications. In *Optical Engineering and Applications, Proceedings of SPIE*, vol. 6669. San Diego, États-Unis.
- Bony A., Bringier B., Khoudeir M. (2013). Tridimensional reconstruction by photometric stereo with near spot light sources. In *Proceedings of the 21st European Signal Processing Conference*. Marrakech, Maroc.
- Boyd S., Parikh N., Chu E., Peleato B., Eckstein J. (2011). Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends in Machine Learning*, vol. 3, n° 1, p. 1–122.
- Ciortan I., Pintus R., Marchioro G., Daffara C., Giachetti A., Gobbetti E. (2016). A Practical Reflectance Transformation Imaging Pipeline for Surface Characterization in Cultural Heritage. In *Proceedings of the 14th Eurographics Workshop on Graphics and Cultural Heritage*. Gênes, Italie.
- Collins T., Bartoli A. (2012). 3D Reconstruction in Laparoscopy with Close-Range Photometric Stereo. In *Proceedings of the 15th International Conference on Medical Imaging and Computer Assisted Intervention*, p. 634–642. Nice, France.
- Durix B., Quéau Y., Lucas T., Boumaza J., Durou J.-D., Lauze F. (2016). Étalonnage de sources lumineuses de type LED. In *Actes du 20^{ème} Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle*. Clermont-Ferrand, France.
- Durou J.-D., Aujol J.-F., Courteille F. (2009). Integrating the Normal Field of a Surface in the Presence of Discontinuities. In *Proceedings of the 7th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition, Lecture Notes in Computer Science*, vol. 5681, p. 261–273. Bonn, Allemagne.
- Gardner I. C. (1947). Validity of the cosine-fourth-power law of illumination. *Journal of Research of the National Bureau of Standards*, vol. 39, p. 213–219.
- Giachetti A., Daffara C., Reghelin E., C. Gobbetti, Pintus R. (2015). Light calibration and quality assessment methods for Reflectance Transformation Imaging applied to artworks'

- analysis. In *Optics for Arts, Architecture, and Archaeology V, Proceedings of SPIE*, vol. 9527. Munich, Allemagne.
- Hara K., Nishino K., Ikeuchi K. (2005). Light source position and reflectance estimation from a single view without the distant illumination assumption. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, n° 4, p. 493–505.
- Hernández C., Vogiatzis G., Brostow G. J., Stenger B., Cipolla R. (2007). Non-rigid Photometric Stereo with Colored Lights. In *Proceedings of the 11th IEEE International Conference on Computer Vision*. Rio de Janeiro, Brésil.
- Hinkley D. V. (1969). On the ratio of two correlated normal random variables. *Biometrika*, vol. 56, n° 3, p. 635–639.
- Horn B. K. P., Brooks M. J. (Eds.). (1989). *Shape from Shading*. MIT Press.
- Huang X., Walton M., Bearman G., Cossairt O. (2015). Near light correction for image re-lighting and 3D shape recovery. In *Proceedings of the International Congress on Digital Heritage*, vol. 1, p. 215–222. Grenade, Espagne.
- McGunnigle G., Chantler M. J. (2003). Resolving handwriting from background printing using photometric stereo. *Pattern Recognition*, vol. 36, n° 8, p. 1869–1879.
- Mecca R., Quéau Y., Logothetis F., Cipolla R. (2016). A Single-Lobe Photometric Stereo Approach for Heterogeneous Material. *SIAM Journal on Imaging Sciences*, vol. 9, n° 4, p. 1858–1888.
- Mecca R., Wetzler A., Bruckstein A. M., Kimmel R. (2014). Near Field Photometric Stereo with Point Light Sources. *SIAM Journal on Imaging Sciences*, vol. 7, n° 4, p. 2732–2770.
- Or-el R., Rosman G., Wetzler A., Kimmel R., Bruckstein A. M. (2015). RGBD-Fusion: Real-Time High Precision Depth Recovery. In *Proceedings of the 28th IEEE Conference on Computer Vision and Pattern Recognition*. Boston, États-Unis.
- Papadimitri T., Favaro P. (2014). Uncalibrated Near-Light Photometric Stereo. In *Proceedings of the 25th British Machine Vision Conference*. Nottingham, Royaume-Uni.
- Park J., Sinha S. N., Matsushita Y., Tai Y.-W., Kweon I. S. (2014). Calibrating a Non-isotropic Near Point Light Source Using a Plane. In *Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition*, p. 2267–2274. Columbus, États-Unis.
- Pintus R., Ciortan I., Giachetti A., Gobbetti E. (2016). Practical Free-form RTI Acquisition with Local Spot Lights. In *Smart Tools and Applications for Graphics, Eurographics Italian Chapter Conference*. Gênes, Italie.
- Powell M. W., Sarkar S., Goldgof D. (2001). A simple strategy for calibrating the geometry of light sources. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, n° 9, p. 1022–1027.
- Quéau Y., Durix B., Lucas T., Boumaza J., Durou J.-D., Lauze F. (2016). Fusion de données RVB-D par stéréophotométrie colorée. In *Actes du 20^{ème} Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle*. Clermont-Ferrand, France.
- Quéau Y., Mecca R., Durou J.-D. (2016). Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach. In *Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, États-Unis.

- Quéau Y., Mecca R., Durou J.-D., Descombes X. (2017). Photometric Stereo with Only Two Images: A Theoretical Study and Numerical Resolution. *Image and Vision Computing*, vol. 57, p. 175–191.
- Quéau Y., Wu T., Lauze F., Durou J.-D., Cremers D. (2017). A Non-Convex Variational Approach to Photometric Stereo under Inaccurate Lighting. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, États-Unis.
- Shen H.-L., Cheng Y. (2011). Calibrating light sources by using a planar mirror. *Journal of Electronic Imaging*, vol. 20, n° 1.
- Sun J., Smith M., Smith L., Farooq A. (2013). Sampling Light Field for Photometric Stereo. *International Journal of Computer Theory and Engineering*, vol. 5, n° 1, p. 14–18.
- Takai T., Maki A., Niinuma K., Matsuyama T. (2009). Difference sphere: An approach to near light source estimation. *Computer Vision and Image Understanding*, vol. 113, n° 9, p. 966–978.
- Woodham R. J. (1980). Photometric Method for Determining Surface Orientation from Multiple Images. *Optical Engineering*, vol. 19, n° 1, p. 139–144.
- Xie L., Song Z., Jiao G., Huang X., Jia K. (2015). A practical means for calibrating an LED-based photometric stereo system. *Optics and Lasers in Engineering*, vol. 64, p. 42–50.
- Xie W., Dai C., Wang C. C. L. (2015). Photometric Stereo With Near Point Lighting: A Solution by Mesh Deformation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, États-Unis.
- Yang J., Ye X., Li K., Hou C., Wang Y. (2014). Color-Guided Depth Recovery From RGB-D Data Using an Adaptive Autoregressive Model. *IEEE Transactions on Image Processing*, vol. 23, n° 8, p. 3443–3458.
- Yeh C.-K., Matsuda N., Huang X., Li F., Walton M., Cossairt O. (2016). A Streamlined Photometric Stereo Framework for Cultural Heritage. In *Proceedings of the 14th European Conference on Computer Vision*, p. 738–752. Amsterdam, Pays-Bas.